

Generalized Adjoint for Physical Processes with Parameterized Discontinuities. Part VI: Minimization Problems in Multidimensional Space

QIN XU

Naval Research Laboratory, Monterey, California

JIDONG GAO

CAPS, University of Oklahoma, Norman, Oklahoma

(Manuscript received 22 December 1997, in final form 18 May 1998)

ABSTRACT

Generalized adjoint with modified discretization and generalized coarse-grain adjoint are derived for a vector system of equations that contains parameterized on/off switches. With vector examples, it is shown that the conventional adjoint minimization may have a convergence problem in multidimensional space. The problem can be solved by the generalized adjoint with modified discretization or by the generalized coarse-grain adjoint without modifying the traditional discretization in the forward model.

1. Introduction

Many physical processes in atmospheric models involve parameterized time discontinuities (on/off switches) that do not satisfy the well-known Lipschitz condition, the condition that ensures the existence, uniqueness, and continuous dependence of the solution on its initial condition in the classical sense. In order to derive the tangent linear and adjoint equations for such a model, it is necessary to extend the classic concept of solution based on the modern mathematical theory on generalized functions (see appendix to chap. 6, Courant and Hilbert 1962). Such an extension was made recently by Xu (1996a,b, 1997a) for time continuous models.

When the generalized adjoint formulations derived for time continuous models are applied to time discrete models, new problems are caused by on/off switches triggered at discrete time levels by threshold conditions. In this case, as shown in Xu (1997b), the discrete solution is not continuously dependent on the initial state, so the cost function contains zigzag discontinuities that cause a problem in the tangent linearization and adjoint minimization, but the problem can be solved by modifying the traditional time discretization with the switch time determined by interpolation as a continuous function of the initial state. The problem can also be avoided by considering the nonlocal coarse-grain geometry of

the cost function, as proposed by Xu et al. (1998, henceforth referred to as XGG98). The generalized coarse-grain tangent linearization and adjoint provide an alternate approach without modifying the traditional time discretization but require the integration time step be sufficiently small.

For the scalar examples in Xu (1997b) and XGG98, there were only two directions for the gradient of the cost function in one-dimensional space, so the minimum of the cost function could be approached by an iterative procedure as long as the sign of the gradient was correctly estimated. Thus, although the conventional adjoint is inaccurate and not suitable for sensitivity studies, the inaccuracy may not cause a serious problem for one-dimensional minimization. This, however, is not the case for minimization in multiple dimensions. The related problems will be examined with vector examples in this paper. The two approaches proposed in Xu (1997b) and XGG98 will be used to derive the generalized adjoint and coarse-grain adjoint for the vector examples.

The paper is organized as follows. The analytical model system and its adjoint are given in the next section. The analytical model is discretized in two ways by using the traditional and modified discretization schemes in section 3. The discrete generalized tangent linear model and the coarse-grain tangent linear model are derived and compared with the conventional tangent linear model in section 4. Discrete adjoints are derived in section 5. Three different minimization procedures are tested and compared in section 6. The results are summarized with conclusions in section 7.

Corresponding author address: Dr. Qin Xu, Naval Research Laboratory, 7 Grace Hopper Drive, Monterey, CA 93943-5502.
E-mail: xuq@nrlmry.navy.mil

2. Model equations

In this paper we use the following modified model of Lorenz (1963):

$$\begin{aligned} d_t x &= -px + py, \\ d_t y &= rx - zx - x, \\ d_t z &= xy - az, \end{aligned} \quad (2.1)$$

where $r = r_0 + r_1 H(c)$ is the modified Rayleigh number that contains a jump controlled by the threshold condition of $c \equiv y - y_c > 0$, and $H(\cdot)$ is the Heaviside unit-step function (see Courant and Hilbert 1962, 622); p and a are related to the Prandtl number and the aspect ratio of geometry (in the original model), respectively. The parameter values used in this paper are $p = 10$, $a = 8/3$, $r_0 = 28$, $r_1 = -18$, and $y_c = -4.5$. According to Lorenz (1963) and Sparrow (1982), for fixed $p = 10$ and $a = 8/3$, the solution changes regime from chaotic to oscillatory asymptotic stable when r decreases from $r_0 = 28$ to $r_0 + r_1 = 10$. With the above parameter settings, the regime changes are controlled by the threshold condition.

To derive the generalized tangent linear and adjoint operators, it is convenient to rewrite (2.1) into the following vector form:

$$d_t \mathbf{x} = \mathbf{F} + H(c)\mathbf{G}, \quad (2.2)$$

where $\mathbf{x} = (x, y, z)^T = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$, $\mathbf{F} = \mathbf{F}(\mathbf{x}) = (-px + py)\mathbf{i} + (r_0x - zx - x)\mathbf{j} + (xy - az)\mathbf{k}$, and $\mathbf{G} = \mathbf{G}(\mathbf{x}) = r_1x\mathbf{j}$. Here, \mathbf{i} , \mathbf{j} , and \mathbf{k} are unit vectors in the x , y , and z directions, respectively. Using (3.2) and (4.2) of X96b with $d_t c_- = d_t y(\tau_-) = \mathbf{j}^T \mathbf{F}|_\tau$ and $\nabla c = \mathbf{j}$, one can derive from (2.2) the following generalized tangent linear and adjoint operators in the vicinity of an on or off switch (at $t = \tau$):

$$L = d_t - \mathbf{P} - H(y - y_c)\mathbf{Q} - H'(t - \tau)\mathbf{S}, \quad (2.3)$$

$$\begin{aligned} L^* &= -d_t - \mathbf{P}^T - H(y - y_c)\mathbf{Q}^T \\ &\quad - H'(t - \tau)\mathbf{S}^T, \end{aligned} \quad (2.4)$$

where $\mathbf{P} = \mathbf{P}(\mathbf{x}) = \nabla \mathbf{F} = [(\mathbf{i}\partial_x + \mathbf{j}\partial_y + \mathbf{k}\partial_z)\mathbf{F}]^T$, $\mathbf{Q} = \mathbf{Q}(\mathbf{x}) = \nabla \mathbf{G} = r_1\mathbf{j}\mathbf{j}^T$, $\mathbf{S} = \mathbf{S}(\mathbf{x}) = \mathbf{G}[(\nabla c)^T]/|\mathbf{j} \cdot \mathbf{F}| = r_1x\mathbf{j}\mathbf{j}^T/|\mathbf{j}^T \mathbf{F}|$, and $H'(\cdot)$ is the unit delta function, the derivative of $H(\cdot)$.

3. Discrete forward models—FM0 and FM1

When the predictor–corrector scheme (Lorenz 1963) is used to discretize (2.1), the switch time can be determined in two ways, similar to the one-dimensional examples in XGG98. The first way uses the traditional discretization, which gives the following discrete model, denoted by FM0:

$$\tilde{\mathbf{x}}_n = \mathbf{x}_{n-1} + [\mathbf{F}(\mathbf{x}_{n-1}) + H(y_{n-1} - y_c)\mathbf{G}(\mathbf{x}_{n-1})]\Delta t$$

and

$$\begin{aligned} \mathbf{x}_n &= \mathbf{x}_{n-1} + \{\mathbf{F}(\mathbf{x}_{n-1}) + \mathbf{F}(\tilde{\mathbf{x}}_n) \\ &\quad + H(y_{n-1} - y_c)[\mathbf{G}(\mathbf{x}_{n-1}) + \mathbf{G}(\tilde{\mathbf{x}}_n)]\}\Delta t/2, \end{aligned} \quad (3.1)$$

where $\mathbf{F}(\cdot)$ and $\mathbf{G}(\cdot)$ are defined as in (2.2) and $\Delta t = T/N$ is the time step. The first formula in (3.1) is the predictor based on the Euler forward scheme, and the second formula is the corrector based on the trapezoidal implicit scheme. Here, as in (3.1) of Xu (1997b) or (2.4) of XGG98, an on (or off) switch is triggered in (3.1) at the discrete time level immediately after the threshold condition is (or is not) exceeded.

The second way uses a modified discretization, and the resulting discrete model is denoted by FM1. For a nonswitch time step, FM1 is the same as FM0. When an on (or off) switch is detected, say, at the m th step by $y_m \geq y_c > y_{m-1}$ (or $y_m < y_c \leq y_{m-1}$), the switch time is determined in FM1 by

$$y_c = y_{m-1} + \mathbf{j}^T \mathbf{F}(\mathbf{x}_{m-1})\Delta\tau, \quad (3.2)$$

where $0 \leq \Delta\tau < \Delta t$ and $\tau = (m-1)\Delta t + \Delta\tau$ is the interpolated switch time. Since the switch in FM1 is triggered at the intermediate time τ (earlier than the switch time $m\Delta t$ in FM0), a new term should be added to (3.1) for the m th time step:

$$\begin{aligned} \tilde{\mathbf{x}}_m &= \mathbf{x}_{m-1} + [\mathbf{F}(\mathbf{x}_{m-1}) + H(y_{m-1} - y_c)\mathbf{G}(\mathbf{x}_{m-1})]\Delta t \\ &\quad \pm \mathbf{G}(\mathbf{x}_{m-1})(\Delta t - \Delta\tau) \end{aligned}$$

and

$$\begin{aligned} \mathbf{x}_m &= \mathbf{x}_{m-1} + [\mathbf{F}(\mathbf{x}_{m-1}) + \mathbf{F}(\tilde{\mathbf{x}}_m)]\Delta t/2 \\ &\quad + H(y_{m-1} - y_c)[\mathbf{G}(\mathbf{x}_{m-1}) + \mathbf{G}(\tilde{\mathbf{x}}_m)]\Delta t/2 \\ &\quad \pm [\mathbf{G}(\mathbf{x}_{m-1}) + \mathbf{G}(\tilde{\mathbf{x}}_m)](\Delta t - \Delta\tau)/2, \end{aligned} \quad (3.3)$$

where \pm signs take plus for an on switch with $H(y_{m-1} - y_c) = 0$ and minus for an off switch with $H(y_{m-1} - y_c) = 1$.

Shown in Fig. 1 is an example of numerical solution obtained by integrating FM1 (solid) from $t = 0$ to $t = T = 200$ with $\Delta t = 0.01$ and $\mathbf{x}_0 = (-6.36, -8.27, 21.2)^T$. During the period of integration, an on switch is triggered at about $t = 37$ when y exceeds the threshold value of $y_c = -4.6$. Since the time step is very small ($\Delta t = 0.01$), the numerical solution obtained by integrating FM0 with the same parameter settings gives a visually almost-identical curve to that in Fig. 1. These two nearly identical numerical solutions will be used as the reference solutions for the verification of the tangent linear solutions in the next section. The numerical solution obtained from FM1 will be used as the observed true state for the cost function defined in section 4.

4. Discrete tangent linear models—CTLMO, GTLM1, and GCTLM

As in XGG98, we denote by CTLMO the conventional tangent linear model derived from FM0, by GTLM1 we denote the generalized tangent linear model derived

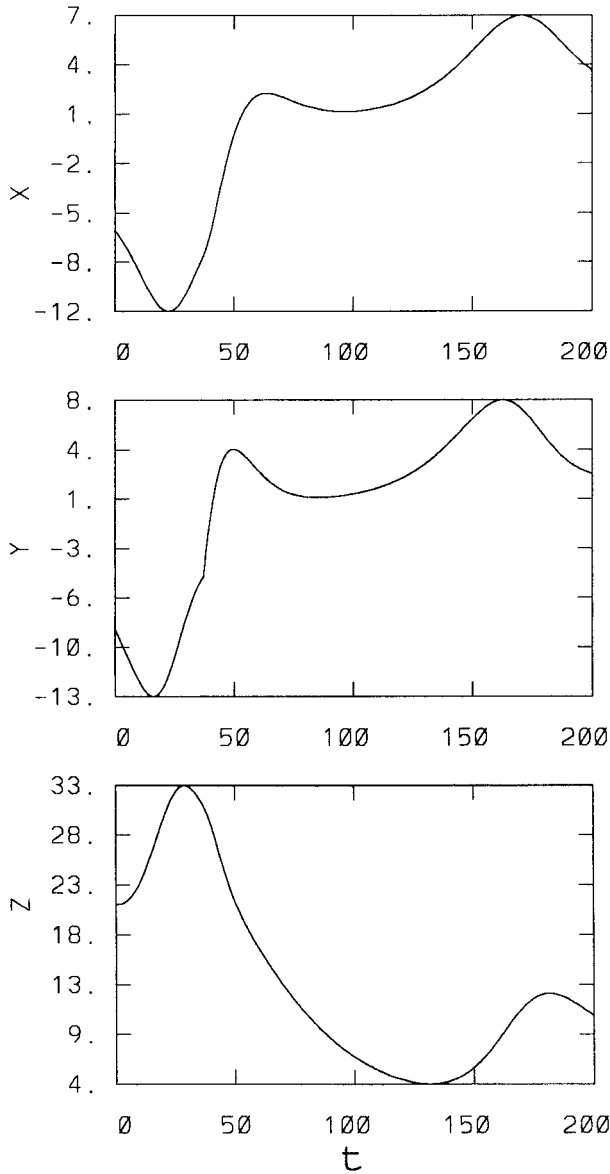


FIG. 1. The three components of the numerical solution obtained by integrating FM1 from $t = 0$ to $t = T = 200$ with $\Delta t = 0.01$ and $\mathbf{x}_0 = (-6.36, -8.27, 21.2)^T$. An on switch is triggered at $t = 0.37$ when y exceeds the threshold value of $y_c = -4.6$.

from FM1, and by GCTLM we denote the generalized coarse-grain tangent linearization of FM0. The conventional tangent linearization ignores the switch time perturbations (or jumps), so $\delta(3.1)$ gives the following conventional discrete tangent linear model (CTLMO):

$$\begin{aligned}\delta\tilde{\mathbf{x}}_n &= \delta\mathbf{x}_{n-1} + [\mathbf{P}(\mathbf{x}_{n-1}) + H(y_{n-1} - y_c)\mathbf{Q}(\mathbf{x}_{n-1})]\delta\mathbf{x}_{n-1}\Delta t, \\ \delta\mathbf{x}_n &= \delta\mathbf{x}_{n-1} + [\mathbf{P}(\mathbf{x}_{n-1})\delta\mathbf{x}_{n-1} + \mathbf{P}(\tilde{\mathbf{x}}_n)\delta\tilde{\mathbf{x}}_n]\Delta t/2 \\ &\quad + H(y_{n-1} - y_c)[\mathbf{Q}(\mathbf{x}_{n-1})\delta\mathbf{x}_{n-1} + \mathbf{Q}(\tilde{\mathbf{x}}_n)\delta\tilde{\mathbf{x}}_n]\Delta t/2,\end{aligned}\quad (4.1)$$

where $\mathbf{P}(\cdot)$ and $\mathbf{Q}(\cdot)$ are defined as in (2.3).

The generalized tangent linearization considers the perturbation of $\Delta\tau$ in (3.3), and $\delta\Delta\tau$ can be determined from $\delta(3.2)$:

$$\begin{aligned}0 &= \delta y_{m-1} + \delta[\mathbf{j}^T\mathbf{F}(\mathbf{x}_{m-1})]\Delta\tau + \mathbf{j}^T\mathbf{F}(\mathbf{x}_{m-1})\delta\Delta\tau \\ &= \delta y_{m-1} + \mathbf{j}^T\mathbf{P}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1}\Delta\tau + \mathbf{j}^T\mathbf{F}(\mathbf{x}_{m-1})\delta\Delta\tau\end{aligned}$$

or

$$\delta\Delta\tau = -[\mathbf{j}^T + \mathbf{j}^T\mathbf{P}(\mathbf{x}_{m-1})\Delta\tau]\delta\mathbf{x}_{m-1}/[\mathbf{j}^T\mathbf{F}(\mathbf{x}_{m-1})]. \quad (4.2)$$

Substituting (4.2) into $\delta(3.3)$ gives GTLM1 for the m th time step at which a switch is triggered:

$$\begin{aligned}\delta\tilde{\mathbf{x}}_m &= \delta\mathbf{x}_{m-1} + [\mathbf{P}(\mathbf{x}_{m-1}) + H(y_{m-1} - y_c)\mathbf{Q}(\mathbf{x}_{m-1})] \\ &\quad \times \delta\mathbf{x}_{m-1}\Delta t \pm \mathbf{Q}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1}(\Delta t - \Delta\tau) \\ &\quad + \mathbf{S}(\mathbf{x}_{m-1})[\mathbf{I} + \Delta\tau\mathbf{P}(\mathbf{x}_{m-1})]\delta\mathbf{x}_{m-1}, \\ \delta\mathbf{x}_m &= \delta\mathbf{x}_{m-1} + [\mathbf{P}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1} + \mathbf{P}(\tilde{\mathbf{x}}_m)\delta\tilde{\mathbf{x}}_m]\Delta t/2 \\ &\quad + H(y_{m-1} - y_c)[\mathbf{Q}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1} + \mathbf{Q}(\tilde{\mathbf{x}}_m)\delta\tilde{\mathbf{x}}_m]\Delta t/2 \\ &\quad \pm [\mathbf{Q}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1} + \mathbf{Q}(\tilde{\mathbf{x}}_m)\delta\tilde{\mathbf{x}}_m](\Delta t - \Delta\tau)/2 \\ &\quad + \beta\mathbf{S}(\mathbf{x}_{m-1})[\mathbf{I} + \Delta\tau\mathbf{P}(\mathbf{x}_{m-1})]\delta\mathbf{x}_{m-1},\end{aligned}\quad (4.3)$$

where $\mathbf{S}(\cdot)$ is defined as in (2.3), $\beta = (1 + \tilde{\mathbf{x}}_m/\mathbf{x}_{m-1})2$, and the rule for \pm signs is as in (3.3). For a nonswitch time step, GTLM1 is the same as CTLMO in (4.1).

For the m th time step at which a switch is triggered, GCTLM can be derived by integrating the analytical tangent linear model [see (2.3)]:

$$\begin{aligned}\delta\tilde{\mathbf{x}}_m &= \delta\mathbf{x}_{m-1} + [\mathbf{P}(\mathbf{x}_{m-1}) + H(y_{m-1} - y_c)\mathbf{Q}(\mathbf{x}_{m-1})]\delta\mathbf{x}_{m-1}\Delta t \\ &\quad \pm \mathbf{Q}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1}(\Delta t - \Delta\tau) + \mathbf{S}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1}, \\ \delta\mathbf{x}_m &= \delta\mathbf{x}_{m-1} + [\mathbf{P}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1} + \mathbf{P}(\tilde{\mathbf{x}}_m)\delta\tilde{\mathbf{x}}_m]\Delta t/2 \\ &\quad + H(y_{m-1} - y_c)[\mathbf{Q}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1} + \mathbf{Q}(\tilde{\mathbf{x}}_m)\delta\tilde{\mathbf{x}}_m]\Delta t/2 \\ &\quad \pm [\mathbf{Q}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1} + \mathbf{Q}(\tilde{\mathbf{x}}_m)\delta\tilde{\mathbf{x}}_m](\Delta t - \Delta\tau)/2 \\ &\quad + \mathbf{S}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1},\end{aligned}\quad (4.4)$$

where \pm signs are as in (4.3) or (3.3), and the impact of the delta function is evaluated by

$$\int_{(m-1)\Delta t}^{m\Delta t} H'(t - \tau)\mathbf{S}\delta\mathbf{x} dt = \mathbf{S}(\tau_-)\delta\mathbf{x}(\tau_-) \approx \mathbf{S}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1}.$$

Note that $\delta\Delta\tau = -\mathbf{j}^T\delta\mathbf{x}_{m-1}/[\mathbf{j}^T\mathbf{F}(\mathbf{x}_{m-1})] + O(\Delta\tau|\delta\mathbf{x}_{m-1}|) \approx -\mathbf{j}^T\delta\mathbf{x}_{m-1}/[\mathbf{j}^T\mathbf{F}(\mathbf{x}_{m-1})]$ in (4.2). When this approximation is used in the derivation of (4.3), the result reduces to (4.4). To the same order of accuracy, (4.4) can further reduce to

$$\begin{aligned}\delta\tilde{\mathbf{x}}_m &= \delta\mathbf{x}_{m-1} + [\mathbf{P}(\mathbf{x}_{m-1}) + H(y_{m-1} - y_c)\mathbf{Q}(\mathbf{x}_{m-1})]\delta\mathbf{x}_{m-1}\Delta t \\ &\quad + \mathbf{S}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1}, \\ \delta\mathbf{x}_m &= \delta\mathbf{x}_{m-1} + [\mathbf{P}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1} + \mathbf{P}(\tilde{\mathbf{x}}_m)\delta\tilde{\mathbf{x}}_m]\Delta t/2 \\ &\quad + H(y_{m-1} - y_c)[\mathbf{Q}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1} + \mathbf{Q}(\tilde{\mathbf{x}}_m)\delta\tilde{\mathbf{x}}_m]\Delta t/2 \\ &\quad + \mathbf{S}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1},\end{aligned}\quad (4.5)$$

or even simply to

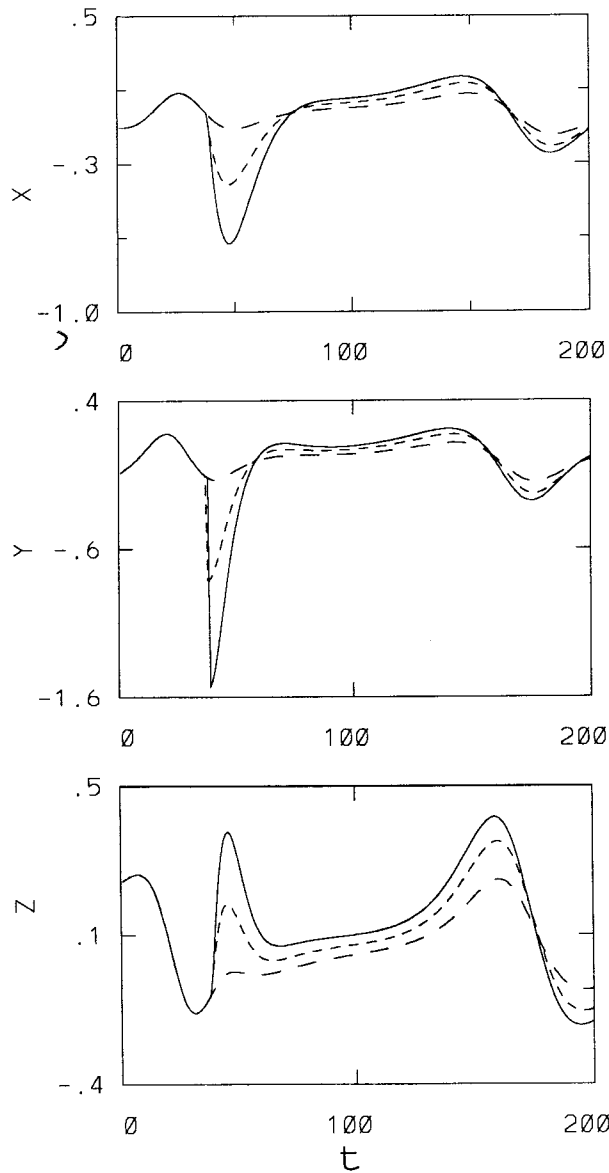


FIG. 2. CTLM0 solution (long-dashed) and GCTLM solution (dashed) obtained from (4.5) in comparison with the nonlinear perturbation (solid) obtained from FM0. The reference initial state \mathbf{x}_0 is the same as in Fig. 1, and the initial perturbation is $\delta\mathbf{x}_0 = 0.01\mathbf{x}_0$. Since $\Delta t = 0.01$ and $|\mathbf{F}| \approx O(10)$, $O(\mathbf{F}\Delta t) < O(\delta\mathbf{x}_0) < 1$ is satisfied.

$$\delta\mathbf{x}_m = \delta\mathbf{x}_{m-1} + \mathbf{S}(\mathbf{x}_{m-1})\delta\mathbf{x}_{m-1}. \quad (4.6)$$

Clearly, to the same order of accuracy, GCTLM can have different forms. GCTLM in (4.4) is most close to GTLM1; GCTLM in (4.5) contains all the conventional terms in CTLM0; and GCTLM in (4.6) has the simplest form. Like the GTLM1 operator in (4.3), the three GCTLM operators in (4.4)–(4.6) all converge to the analytical TLM operator in (2.3) as $\Delta t \rightarrow 0$. The CTLM operator in (4.1), however, does not converge to the analytical TLM operator in (2.3). For $\Delta\tau \leq \Delta t \ll 1$, the difference between GCTLM and GTLM1 is much

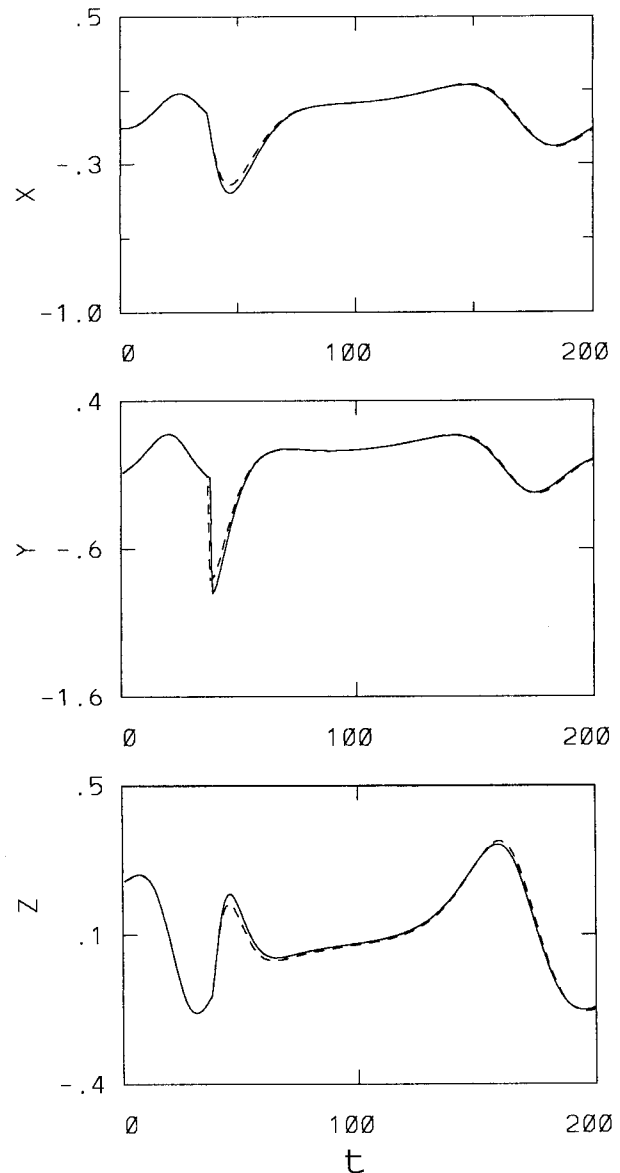


FIG. 3. As in Fig. 2, but for the GTLM1 solution (dashed) in comparison with the nonlinear perturbation (solid) obtained from FM1 for a very small initial perturbation $\delta\mathbf{x}_0 = 0.01\mathbf{x}_0$ satisfying $O(\delta\mathbf{x}_0) \leq O(\mathbf{F}\Delta t) < 1$.

smaller [by a factor of $O(\Delta t)$] than the difference between CTLM0 and GTLM1. As shown in Fig. 2, the CTLM0 solution (long-dashed) does not follow the nonlinear perturbation (solid) obtained from FM0 after the on switch, but the GCTLM solution (dashed) obtained from (4.5) can follow the nonlinear perturbation. Thus, as explained in XGG98, GCTLM can be used as a coarse-grain tangent linear model of FM0 as long as Δt is sufficiently small. When Δt is not sufficiently small, GCTLM may not be used as a coarse-grain tangent linear model of FM0, but it is still a valid approximation of GTLM1. As shown in Fig. 3, the GTLM1 solution

(dashed) closely follows the nonlinear perturbation (solid) obtained from FM1. The GCTL solution (dashed in Fig. 2) also follows the nonlinear perturbation obtained from FM1 (solid in Fig. 3) but not as closely as the GTLM1 solution.

5. Discrete adjoint and gradient check

The discrete cost function can be defined by

$$J_d(\mathbf{x}_0) = \sum_{n=0}^N |\mathbf{D}_n|^2 \Delta t = \sum_{n=0}^N |\mathbf{x}_n - \mathbf{x}_{\text{obsn}}|^2 \Delta t, \quad (5.1)$$

where $\mathbf{x}_n = \mathbf{x}_n(\mathbf{x}_0)$ is a function of initial state constrained by the discrete forward model (FM0 or FM1), and \mathbf{x}_{obsn} denotes the observed value of \mathbf{x} at the n th time level. For the numerical experiments in this paper, \mathbf{x}_{obsn} is given by the FM1 solution with $\mathbf{x}_0 = (x_0, y_0, z_0)^T = (-6.36, -8.27, 21.2)^T$, that is, the solid curve in Fig. 1. Note from (5.1) that discontinuities caused by $\mathbf{x}_n(\mathbf{x}_0)$ in the cost function are independent of observations. Thus, including random observational errors in \mathbf{x}_{obsn} will not affect the nature of the problem examined in this paper. This is verified by our numerical experiments with imperfect observations (not shown), so observational errors are not an issue in this paper. When \mathbf{x}_n in (5.1) is constrained by FM1, the minimum of the cost function is at $\mathbf{x}_0 = (-6.36, -8.27, 21.2)^T$, which is the initial state for the solution in Fig. 1. Because the FM0 solution is very close to the FM1 solution (see Fig. 1), when \mathbf{x}_n is constrained by FM0, the minimum of the cost function is almost exactly at the same \mathbf{x}_0 point of the initial state for the solution in Fig. 1.

In association with the cost function in (5.1), CTLM0 in (4.1) gives the following conventional discrete adjoint model (CAJM0):

$$\begin{aligned} \delta \tilde{\mathbf{x}}_n^* &= [\mathbf{P}^T(\tilde{\mathbf{x}}_n) + H(y_{n-1} - y_c) \mathbf{Q}^T(\tilde{\mathbf{x}}_n)] \delta \mathbf{x}_n^* \Delta t / 2, \\ \delta \mathbf{x}_{n-1}^* &= \delta \mathbf{x}_n^* + \delta \tilde{\mathbf{x}}_n^* + \mathbf{P}^T(\mathbf{x}_{n-1})(\delta \tilde{\mathbf{x}}_n^* + \delta \mathbf{x}_n^* / 2) \Delta t \\ &\quad + H(y_{n-1} - y_c) \mathbf{Q}^T(\mathbf{x}_{n-1})(\delta \tilde{\mathbf{x}}_n^* + \delta \mathbf{x}_n^* / 2) \Delta t \\ &\quad + 2\mathbf{D}_n \Delta t. \end{aligned} \quad (5.2)$$

For the m th time step at which a switch is triggered, GTLM1 in (4.3) yields the following generalized discrete adjoint model (GAJM1):

$$\begin{aligned} \delta \tilde{\mathbf{x}}_m^* &= [\mathbf{P}^T(\tilde{\mathbf{x}}_m) + t_3 H(y_{m-1} - y_c) \mathbf{Q}^T(\tilde{\mathbf{x}}_m)] \delta \mathbf{x}_m^* \Delta t / 2 \\ &\quad \pm \mathbf{Q}^T(\mathbf{x}_{m-1}) \delta \mathbf{x}_m^* (\Delta t - \Delta \tau) / 2, \\ \delta \mathbf{x}_{m-1}^* &= \delta \mathbf{x}_m^* + \delta \tilde{\mathbf{x}}_m^* + H(y_{m-1} - y_c) \mathbf{Q}^T(\mathbf{x}_{m-1}) \\ &\quad \times (\delta \tilde{\mathbf{x}}_m^* + \delta \mathbf{x}_m^* / 2) \Delta t + \mathbf{P}^T(\mathbf{x}_{m-1})(\delta \tilde{\mathbf{x}}_m^* + \delta \mathbf{x}_m^* / 2) \Delta t \\ &\quad + \mathbf{P}^T(\mathbf{x}_{m-1})(\delta \tilde{\mathbf{x}}_m^* + \delta \mathbf{x}_m^* / 2) \Delta \tau \\ &\quad + [\mathbf{I} + \mathbf{P}^T(\mathbf{x}_{m-1}) \Delta \tau] \mathbf{S}^T(\mathbf{x}_{m-1})(\delta \tilde{\mathbf{x}}_m^* + \beta \delta \mathbf{x}_m^*) \\ &\quad \pm \mathbf{Q}^T(\mathbf{x}_{m-1})(\delta \mathbf{x}_m^* / 2 + \delta \tilde{\mathbf{x}}_m^*)(\Delta t - \Delta \tau) 2\mathbf{D}_m \Delta t. \end{aligned} \quad (5.3)$$

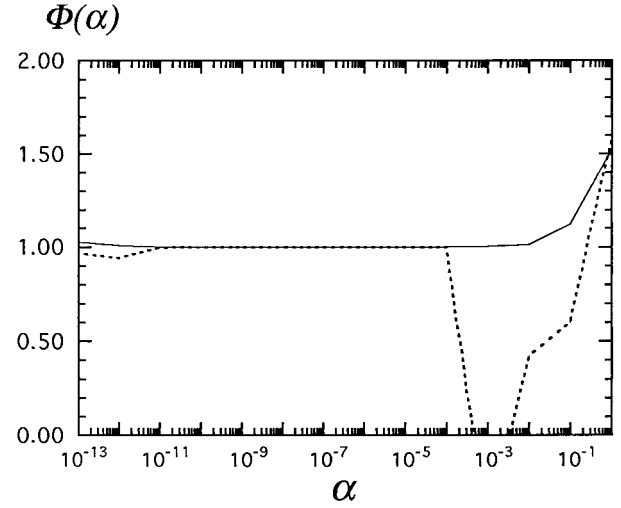


FIG. 4. The $\Phi(\alpha)$ curves with gradient computed by CAJM0 and cost function computed by FM0 at $\mathbf{x}_0 = (-3.86, -8.77, 21.2)^T$ with $\Delta t = 0.01$ (solid) and $\Delta t = 0.0001$ (dotted).

Similarly, for the m th time step at which a switch is triggered, GCTL in (4.5) yields the following discrete adjoint model (GAJM):

$$\begin{aligned} \delta \tilde{\mathbf{x}}_m^* &= [\mathbf{P}^T(\tilde{\mathbf{x}}_m) + H(y_{m-1} - y_c) \mathbf{Q}^T(\tilde{\mathbf{x}}_m)] \delta \mathbf{x}_m^* \Delta t / 2, \\ \delta \mathbf{x}_{m-1}^* &= \delta \mathbf{x}_m^* + \delta \tilde{\mathbf{x}}_m^* + H(y_{m-1} - y_c) \mathbf{Q}^T(\mathbf{x}_{m-1}) \\ &\quad \times (\delta \tilde{\mathbf{x}}_m^* + \delta \mathbf{x}_m^* / 2) \Delta t + \mathbf{P}^T(\mathbf{x}_{m-1})(\delta \tilde{\mathbf{x}}_m^* + \delta \mathbf{x}_m^* / 2) \Delta t \\ &\quad + \mathbf{S}^T(\mathbf{x}_{m-1})(\delta \tilde{\mathbf{x}}_m^* + \delta \mathbf{x}_m^*) + 2\mathbf{D}_m \Delta t. \end{aligned} \quad (5.4)$$

Over a smooth part of the cost function, the accuracy of the gradient computed by an adjoint model can be checked by the truncated Taylor expansion of the cost function, which leads to the following formula:

$$\begin{aligned} \Phi(\alpha) &= [J_d(\mathbf{x}_0 + \alpha \mathbf{u}) - J_d(\mathbf{x}_0)] / [\alpha \mathbf{u} \cdot \nabla J_d(\mathbf{x}_0)] \\ &= 1 + O(\alpha), \end{aligned} \quad (5.5)$$

where $\mathbf{u} = \nabla J_d(\mathbf{x}_0) / |\nabla J_d(\mathbf{x}_0)|$ and α is a scalar much smaller than unity. When α is small but not as small as the machine round-off error, one should expect that $\Phi(\alpha) = 1 + O(\alpha)$. When the gradient computed by CAJM0 is checked by (5.5), α should be sufficiently small so that the switches do not jump in the FM0 solution. Depending on the reference initial state \mathbf{x}_0 and/or time step Δt used in the integration, the required smallness of α can be very different. The situation is similar to that in Fig. 6 of XGG98, and the problem is caused by the noncontinuous dependence of the FM0 solution on the initial state \mathbf{x}_0 . As shown in Fig. 4, with $\mathbf{x}_0 = (-3.86, -8.77, 21.2)^T$ and $\Delta t = 0.01$, the $\Phi(\alpha)$ curve (solid) is very close to 1 over the wide range of $10^{-12} < \alpha < 10^{-2}$. However, when $\Delta t = 0.0001$, the $\Phi(\alpha)$ curve (dotted) drops quickly far below 1 when α becomes larger than 10^{-4} .

The gradient computed by GAJM1 can be checked

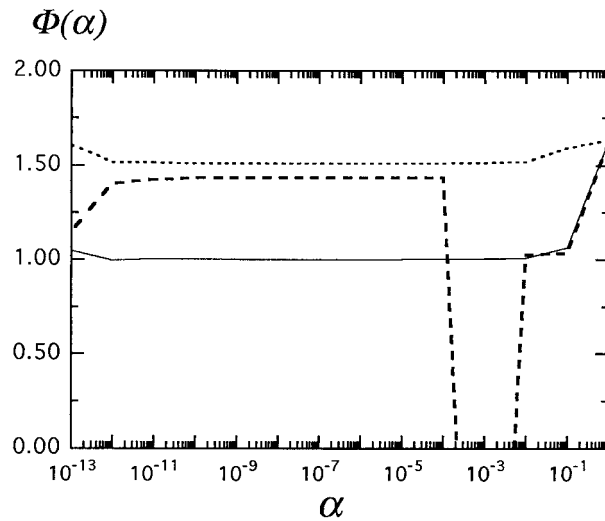


FIG. 5. The $\Phi(\alpha)$ curve with gradient computed by GAJM1 and cost function computed by FM1 (solid) at $\mathbf{x}_0 = (-3.86, -8.77, 21.2)^T$ with $\Delta t = 0.01$ (solid); and $\Phi(\alpha)$ curves with gradient computed by GCAJM and cost function computed by FM0 at $\mathbf{x}_0 = (-3.86, -8.77, 21.2)^T$ with $\Delta t = 0.01$ (dotted) and $\Delta t = 0.0001$ (dashed).

by (5.5) to a great accuracy. As shown in Fig. 5, the $\Phi(\alpha)$ curve (solid) is very close to 1 over the wide range of $10^{-12} < \alpha < 10^{-2}$, and this result is independent of \mathbf{x}_0 and Δt . When the gradient is computed by GCAJM and the cost function is computed by FM0 with $\Delta t = 0.01$, the $\Phi(\alpha)$ curve (dotted) is always above 1 and the gradient check fails. However, when the cost function is computed by FM0 with $\Delta t = 0.0001$, the dotted $\Phi(\alpha)$ curve in Fig. 5 is very close to 1 over the range of $10^{-2} < \alpha < 10^{-1}$. On the other hand, the dotted $\Phi(\alpha)$ curve in Fig. 4 is far below 1 over this range. Thus, when Δt is sufficiently small, the nonlocal coarse-grain gradient can be estimated by GCAJM, and GCAJM can be used as a coarse-grain adjoint model for FM0.

6. Minimization experiments

By using the above-derived three adjoint models, three minimization methods can be designed. The first method, denoted by FM0-CAJM0, uses FM0 for the forward integration and CAJM0 for the backward gradient computation. The second method, denoted by FM1-GAJM1, uses FM1 for the forward integration and GAJM1 for the backward gradient computation. The third method, denoted by FM0-GCAJM, uses FM0 for the forward integration and GCAJM in (5.4) for the backward gradient computation. The memoryless quasi-Newton algorithm (Liu and Nocedal 1988) is used to search for the minimum of the cost function.

Numerical experiments are performed with $\Delta t = 0.01$. The descending procedures are started from the same first guess at $(x_0, y_0, z_0) = (-3.86, -8.77, 21.2)^T$, with $z_0 = z_{\text{obs}} = 21.2$ fixed at the observed value. In this way, the cost function minimum is searched only

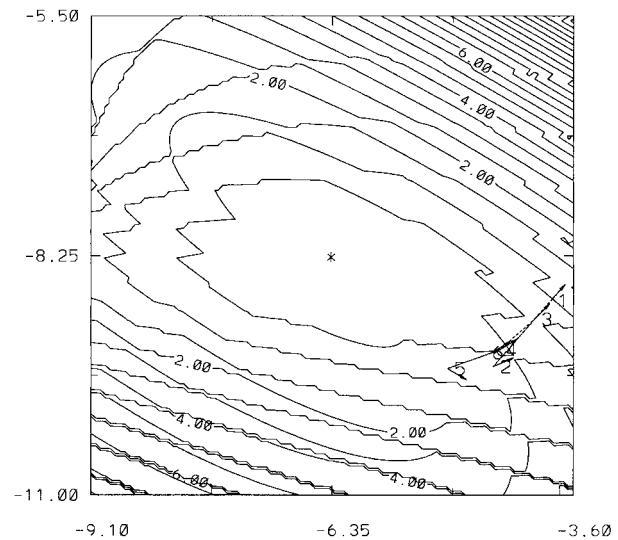


FIG. 6. Costfunction (contours) and descending process (dotted lines) in the two-dimensional space (x_0, y_0) for experiment FM0-CAJM0 with $\Delta t = 0.01$. The first guess is $(x_0, y_0, z_0) = (-3.86, -8.77, 21.2)^T$. The gradient computed at each step (numbered sequentially) is shown by solid arrow. For clarity, only the first eight steps are shown.

in the two-dimensional space (x_0, y_0) , so the descending steps can be easily illustrated. As shown in Fig. 6 for the FM0-CAJM0 method, the descending procedure does not converge to the minimum because the gradients (shown by solid arrows) computed by CAJM0 do not follow the global coarse-grain geometry of the cost function. For clarity, only the first six iterative steps are shown in Fig. 6, though the iteration is performed up to 20 steps. After the sixth step, the solution is trapped in the vicinity of point 6 and there is no improvement in convergence toward the true minimum. The zigzag patterns in the cost function contours manifest discontinuities caused by noncontinuous dependence of the FM0 solution on the initial state. The two-dimensional surface of the cost function can be viewed as a “stadium” having many steps. The gradient computed by CAJM0 follows the step surface locally, but the local geometry can be very different from the global coarse-grain geometry of the cost function. This explains why the solution is trapped in the vicinity of point 6 in Fig. 6. When the modified discretization is used in FM1, the cost function becomes continuous and smooth as shown in Fig. 7. In this case, the gradients can be correctly computed by GAJM1, and the descending procedure converges to the minimum rapidly. Figure 8 shows that the coarse-grain gradients can be correctly computed by GCAJM, so the descending procedure can converge approximately to the minimum for the FM0-GCAJM method, although the cost function has the same discontinuities as in Fig. 6. Figures 9 and 10 show that the cost function and gradient decrease only slightly for the FM0-CAJM0 method, but decrease significantly for the

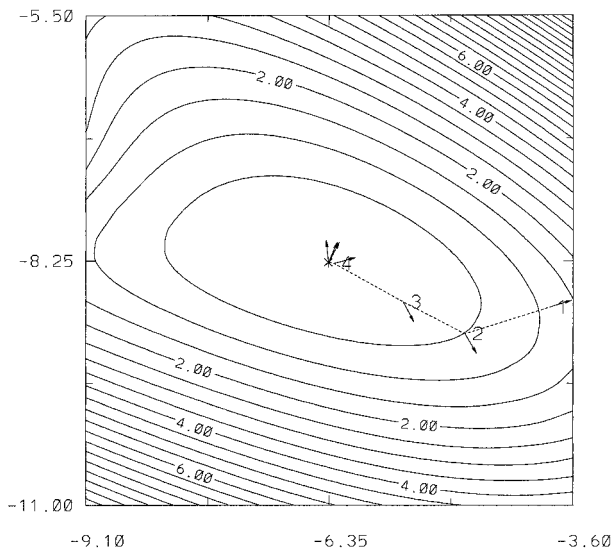


FIG. 7. As in Fig. 6, but for experiment FMI-GAJM1.

FM0-GCAJM method, and decrease to the machine zero for the FM1-GAJM1 method.

Numerical experiments are also performed with different Δt and different first guesses of (x_0, y_0, z_0) . In particular, three different values (0.01, 0.001, 0.0001) are used for Δt , and 40 different first guesses are selected (by equally spaced points) along the four boundaries of the domain in Figs. 6–8. The results of these $(3 \times 3 \times 40)$ experiments can be summarized as follows. 1) With the FM0-CAJM0 method, the descending procedure often does not converge (for 50% of the 40 cases of different first guesses), and the result is very sensitive to the first guess. 2) With the FM1-GAJM1 method, the descending procedure always converges rapidly, the final point (x_0, y_0) is very close to the true minimum

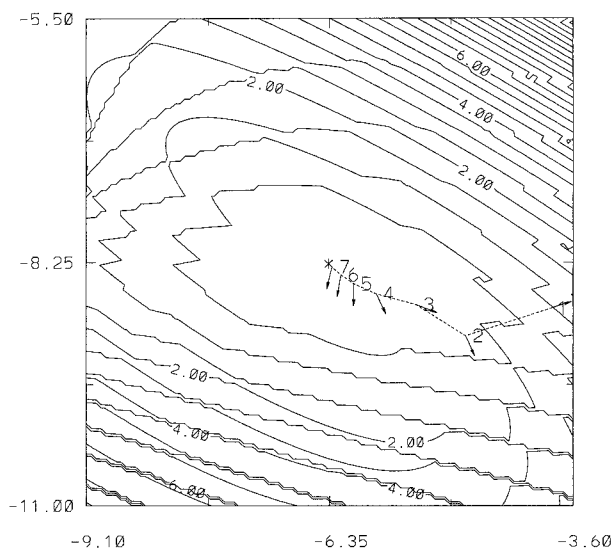
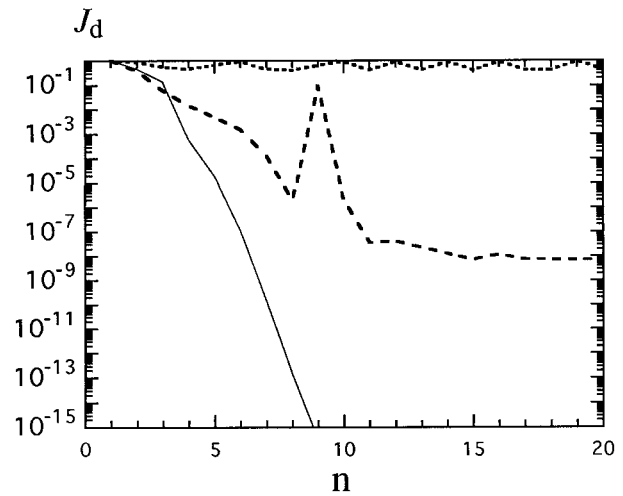


FIG. 8. As in Fig. 6, but for experiment FM0-GCAJM.

FIG. 9. The cost function $J_d(\mathbf{x}_0)$ vs iteration steps in experiments FM0-CAJM0 (dotted), FM0-GCAJM (dashed), and FM1-GAJM1 (solid).

point $(x_0, y_0) = (-3.86, -8.77)$, and the result is virtually independent of the first guess. 3) When the FM0-GCAJM method is used with $\Delta t = 0.01$, the descending procedure converges, and the final point (x_0, y_0) is close to the true minimum point for more than 90% of the 40 cases. The convergence is improved dramatically when Δt is reduced (to 0.001 and 0.0001) in the FM0-GCAJM method. The rms error of the final points (x_0, y_0) (started from 40 different first guesses in each group) with respect to the true minimum point is listed in Table 1 for each method and each selected Δt . Clearly, the rms errors are large for results of FM0-CAJM0, and the error does not decrease with Δt . The rms errors are very small for the results of FM1-GAJM1, and the error decreases rapidly with Δt . For the results of FM1-GAJM1, the rms error decreases rapidly from 0.32 to 0.01 when Δt decreases from 0.01 to 0.001.

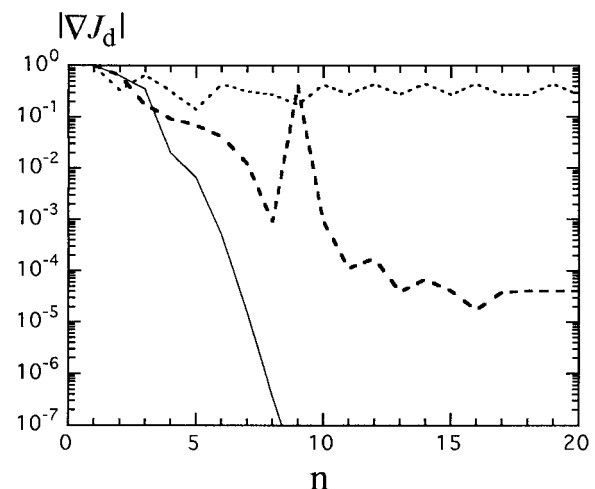
FIG. 10. As in Fig. 9, but for $|\nabla J_d(\mathbf{x}_0)|$ vs iteration steps.

TABLE 1. The rms error of minimum points (x_0, y_0) searched by descending procedures started from 40 different first guesses for each method and given Δt .

	$\Delta t = 0.01$	$\Delta t = 0.001$	$\Delta t = 0.0001$
FM0-CAJMO	1.005	1.088	1.184
FM1-GAJM1	1.52×10^{-2}	0.456×10^{-2}	0.174×10^{-3}
FMO-GCAJM	0.321	1.06×10^{-2}	0.736×10^{-2}

In the above experiments, the minimization procedures are performed in the two-dimensional space of (x_0, y_0) in order to make clear illustrations (see Figs. 6–8). Additional experiments are also performed with different initial guesses of \mathbf{x}_0 and for minimization procedures in the three-dimensional space of the initial state (x_0, y_0, z_0) of the modified Lorenz model. The results remain qualitatively the same.

7. Conclusions

The vector equation system of Lorenz (1963) is modified and used to study how the generalized adjoint theory and analytical formulations in X96b can be applied to time discrete models. As in XGG98, the analytical model system is discretized in two ways by using the traditional and modified discretization schemes, and the resulting discrete models are denoted by FM0 and FM1, respectively. Corresponding to FM0 and FM1, three types of discrete tangent linear models are derived: the conventional tangent linear model (CTLM0) derived from FM0 by ignoring the perturbation of switch time, the generalized tangent linear model (GTLM1) derived from FM1, and the coarse-grain tangent linear model (GCTLM) derived by directly discretizing the analytical tangent linear equations. From these tangent linear models, three discrete adjoint models are derived. The results obtained with vector examples in this paper support the principle results summarized in the conclusion section of XGG98.

Vector examples are used in this paper to illustrate problems in the conventional adjoint minimization and to examine how the problem can be solved by the generalized adjoint with modified discretization or by the coarse-grain adjoint without modifying the traditional discretization in the forward model. The results are summarized as follows.

- 1) The conventional adjoint can compute the local gradient of the zigzag discontinuous cost function constrained by FM0, but the local gradient can be very different from the global coarse-grain gradient. This can cause the conventional adjoint minimization fail to converge.
- 2) The above problem can be solved if FM1 is used as the forward model (in which the switch time is determined by interpolation as a continuous function of the initial state) and the generalized adjoint is used to compute the gradient of the cost function constrained by FM1.
- 3) Without modifying the traditional discretization, the coarse-grain adjoint model can be used to compute the coarse-grain gradient of the zigzag discontinuous cost function constrained by FM0. The convergence of the coarse-grain adjoint minimization can be ensured and improved if a small time step is used for the time integrations of the forward model and backward adjoint model.
- 4) Observational errors are not an issue for the problems examined in this paper, because discontinuities caused by the solution in the cost function are independent of observations. Actually, all the basic findings obtained with perfect observations are confirmed by numerical experiments with imperfect observations (not presented in this paper).
- 5) Although a complete cost function in data assimilation should include a background term, the nature of the problems examined in this and our previous studies is independent of the (neglected) background term.

As shown by Xu (1996a, section 7), when the parameterized discontinuity is fitted by a continuous or smooth function of the control variable, the variation of the switch point is implicitly considered by the variation of the control variable, and this makes the switch suitable for the conventional adjoint method. This type of treatment was used previously by Verlinde and Cotton (1993) and Zupanski and Mesinger (1995). The method is relatively straightforward but requires that the original threshold condition be modified. The generalized adjoint proposed by Xu (1997b) does not change the threshold condition but requires that the traditional discretization be modified (only in the forward model). The generalized coarse-grain adjoint does not change the original forward model but requires that the integration time step be sufficiently small. Each method has certain advantages and disadvantages. Comparisons between these different methods deserve further studies.

Acknowledgments. We are thankful to anonymous reviewers for their comments that improved the presentation of the results. This work was supported by the NSF Grants ATM-9417304 and ATM91-20009 to CIMMS/CAPS at the University of Oklahoma, and by the Office of Naval Research, Program Element 0602435N to the Marine Meteorology Division of the Naval Research Laboratory at Monterey.

REFERENCES

- Courant, R., and D. Hilbert, 1962: *Methods of Mathematical Physics*. Vol. 2. Interscience Publishing, 830 pp.
- Liu, D.C., and J. Nocedal, 1988: On the limited memory BFGS method for large scale optimization. Tech. Rep. NAM03, 26 pp. [Available from Dept. of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL 60208.]
- Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141.
- Sparrow, C., 1982: *The Lorenz Equations*. Springer-Verlag, 269 pp.

- Verlinde, J., and W. R. Cotton, 1993: Fitting microphysical observations of nonsteady convective clouds to a numerical model: An application of the adjoint technique of data assimilation to a kinematic model. *Mon. Wea. Rev.*, **121**, 2776–2793.
- Xu, Q., 1996a: Generalized adjoint for physical processes with parameterized discontinuities. Part I: Basic issues and heuristic examples. *J. Atmos. Sci.*, **53**, 1123–1142.
- , 1996b: Generalized adjoint for physical processes with parameterized discontinuities. Part II: Vector formulations and matching conditions. *J. Atmos. Sci.*, **53**, 1143–1155.
- , 1997a: Generalized adjoint for physical processes with parameterized discontinuities. Part III: Multiple threshold conditions. *J. Atmos. Sci.*, **54**, 2713–2721.
- , 1997b: Generalized adjoint for physical processes with parameterized discontinuities. Part IV: Problems and treatment in time discretization. *J. Atmos. Sci.*, **54**, 2722–2728.
- , J. Gao, and W. Gu, 1998: Generalized adjoint for physical processes with parameterized discontinuities. Part V: Coarse-grain adjoint and problems in gradient check. *J. Atmos. Sci.*, **55**, 2130–2135.
- Zupanski, D., and F. Mesinger, 1995: Four-dimensional data assimilation of precipitation data. *Mon. Wea. Rev.*, **123**, 1112–1127.